ZMK

UNIVERSITÄT HAMBURG

Zentrum für Meeres- und Klimaforschung

The extended OMP Analysis

An analysis package for MATLAB

Version 1

Johannes Karstensen

Institut für Meereskunde der Universität Hamburg

Troplowitzstrasse 7

22529 Hamburg

Germany

email: J.Karstensen@omnet.com

Technical report, 1-99

Institut für Meereskunde

Troplowitzstraße 7
D-22529 Hamburg

Telefon: (040) 428 38-26 05 · Telefax: (040) 560 59 26
http://www.ifm.uni-hamburg.de

The extended OMP Analysis

An analysis package for MATLAB

Version 1

Johannes Karstensen
Institut für Meereskunde der Universität Hamburg
Troplowitzstrasse 7
22529 Hamburg
Germany
email: J.Karstensen@omnet.com

Technical report ; 1 - 99

CONTENTS

1. PREFACE

This report is a hardcopy of the first version of the user guide for the extended OMP analysis. The most current version should be available on the internet at http://www.ifm.uni-hamburg.de/~wwwro/omp_std or send an email to the author J.Karstensen@omnet.com.
If you intend to use the MATLAB package or if you have any questions or comments on OMP analysis, please send them to the *OMP User group* under the above mentioned adress.

2. INTRODUCTION

The knowledge about exact mixing fractions of water masses in the oceans is necessary for various applications and in particular when analysing transient tracer fields or biogeochemical cycling. The distributions of tracers are controlled by a combination of transport processes associated with the oceanic circulation and mixing and by reactive processes associated with the major biogeochemical cycles [3]. To evaluate the distribution of dissolved constituents in sea water one has to resolve effects of mixing and of biogeochemical cycling. OMP analysis is a tool to decompose mixed water masses into fractions of the original water masses (so called source water masses - SWT).

The program package described in this manual is a first try on setting up an easy-to-use OMP analysis. Because the package is far from being complete please let me know if you are using it, at least for mailing you the latest changes and bug-report:

email: J.Karstensen@omnet.com
fax: + 49 40 4123 6631

Water mass mixing analysis was first introduced by [4], while [13] extended it with his "core-layer" method as a tool to resolve information about oceanic circulation. [11] extended the classical mixing triangle, which uses temperature, salinity and mass conservation to determine the mixing fractions from three water masses, into a system of linear mixing equations for more than three water masses by including oxygen and nutrients. This technique allows the determination of mixing fractions on regional scales, where biogeochemical influences on the parameter fields can be neglected. Several authors [10, 7, 12] contributed to the important step of defining the mixing equations as an overdetermined problem where the mixing fractions are determined through optimisation. This technique has become known as Optimum Multiparameter analysis or OMP analysis.

The document is structured as follows: After a short excursion into OMP analysis theory, a general set-up for an appropriate data format is given, and the package files are described. To make the procedure easier to understand, a test dataset from the Indonesian Throughflow region (Indian Ocean Thermocline) is analysed. Additionally, a GUI version of the package is explained.

3. OMP ANALYSIS

3.1 Basics and Classical OMP Analysis

A short description of the theory of OMP analysis is given here, for further details the reader is referred to [7, 12] and [5].

OMP analysis is based on the assumption that mixing is a linear process and affects all parameters in the same way (as this occurs, for example, in eddy turbulence). Starting from this premise OMP analysis tries to invert an oceanic mixing situation with the aid of predefined source water types (SWT), the parameter values found in the "unmixed" state. The observed parameter values which are to be analysed for contributions (mixing fractions) from the original SWT should be linear combinations of the SWT values; this means that the region should be "downstream" from the formation regions of the source water masses and on the spreading path. The SWT contributions or fractions (named x_i) for each data point are obtained by finding the best linear mixing combination in the parameter space defined by, say, temperature, salinity, oxygen, and nutrients which minimises the residuals in a non-negative least squares sense [6].

The solution includes two physically realistic constraints: (a) The contributions from all source water types must sum up to 100%, and (b) all source water type contributions must be non-negative.

As an example we assume a mixing situation of four source water types and five parameters, the linear system of equations building the classical OMP analysis [7, 12] will look like:

$$\begin{array}{rcrcrcrcl}
x_1 T_1 & + & x_2 T_2 & + & x_3 T_3 & + & x_4 T_4 & = & T_{\mathrm{obs}} + R_T \\
x_1 S_1 & + & x_2 S_2 & + & x_3 S_3 & + & x_4 S_4 & = & S_{\mathrm{obs}} + R_S \\
x_1 \mathrm{O}_{2,1} & + & x_2 \mathrm{O}_{2,2} & + & x_3 \mathrm{O}_{2,3} & + & x_4 \mathrm{O}_{2,4} & = & \mathrm{O}_{2,\mathrm{obs}} + R_{\mathrm{O}_2} \\
x_1 \mathrm{PO}_{4,1} & + & x_2 \mathrm{PO}_{4,2} & + & x_3 \mathrm{PO}_{4,3} & + & x_4 \mathrm{PO}_{4,4} & = & \mathrm{PO}_{4,\mathrm{obs}} + R_{\mathrm{PO}_4} \\
x_1 \mathrm{NO}_{3,1} & + & x_2 \mathrm{NO}_{3,2} & + & x_3 \mathrm{NO}_{3,3} & + & x_4 \mathrm{NO}_{3,4} & = & \mathrm{NO}_{3,\mathrm{obs}} + R_{\mathrm{NO}_3} \\
x_1 & + & x_2 & + & x_3 & + & x_4 & = & 1 + R_\Sigma
\end{array}$$

Here the observed values of temperature T_{obs}, salinity S_{obs}, oxygen $\mathrm{O}_{2,\mathrm{obs}}$, phosphate $\mathrm{PO}_{4,\mathrm{obs}}$, and nitrate $\mathrm{NO}_{3,\mathrm{obs}}$ and their respective residuals R define the columns on the right-hand side. The values T_i, S_i, $\mathrm{O}_{2,i}$, $\mathrm{PO}_{4,i}$, and $\mathrm{NO}_{3,i}$ define the fixed parameter values of the four source water types for each parameter. The last row is an expression of the mass conservation. The above system of equations can be written in a short form as:

$$G\, x = d + R$$

where G is the source water type matrix, d is the observational data vector, R is the residual vector of the fit and x is the vector of the source water type fractions we are looking for.

3.2 Normalisation and Weighting of the System

To find a solution to the system of equations, a normalisation (G') of the source water type matrix (G)is necessary in order to make parameters of incommensurable units comparable [12]. This is achieved by normalising the elements in G by the total range of each parameter in G. If j indicates the row index (goes from 1 to $m-1$ parameters) and i the column index (goes from 1 to n unknown), we can obtain the normalised system (G'_{ji}) as:

$$G'_{ji} = (G_{ji} - \overline{G}_j)/\sigma_j$$

Here, $\overline{G}_j$ is the mean source water type of one parameter j calculated as:

$$\overline{G}_j = 1/n \sum_{i=1}^{n} G_{ji}$$

and σ_j is the standard deviation within each row, which is:

$$\sigma_j = \sqrt{1/n \sum_{i=1}^{n} (G_{ji} - \overline{G}_j)^2}$$

The normalised variables are nondimensional, with zero mean and a display variation of order 1.

In order to take care of the differences in parameter accuracy and of the variability in the source region, a weighting procedure is necessary. Weights can be obtained by relating each parameter variance in the source water matrix G to a measure of the variance in the source region. As suggested by *Tomczak and Large* [1989], the largest variance of the parameter in the source region ($\delta_{j\ \max}$) can be used:

$$W_j = \sigma_j^2/\delta_{j\ \max}$$

Weighting procedures discussing covariances between tracers due to variability of the source water types can be found in *Mackas et al.* [1987]. In our example we will use a mean variance in the source region obtained from the MATLAB script POLYVAL.M.

3.3 Extended OMP Analysis

The procedure described above as Basic and Classical OMP Analysis is based on the assumption that observed changes in parameter distributions are exclusively due to mixing and advection. This works well if the analysis extends only over a limited ocean region. If one is more interested in analysing the water mass structure on a larger scale, biogeochemical changes in the parameter field have to be included in the OMP analysis. Introducing 'PO' or 'NO' [2] is one way to achieve this [15, 14]. A different way is by adding another column representing the biogeochemical transfer to the water type matrix [5]. The advantage of doing this is taht the change in the biogeochemical parameter field are obtained directly as an additionally output. In each case one looses on degree of freedom (compared with the "classical OMP"analysis version (section 3.1)) by adding oxygen and nitrate to 'NO' or oxygen and phosphate to 'PO' or adding a new unknown ΔP, to resolve the biogeochemical changes:

$x_1 T_1$	$+\cdots+$	$x_5 T_5$	$+$	0	$+$	0	$+$	0	$=$	T_{obs}	$+$	R_T
$x_1 S_1$	$+\cdots+$	$x_5 S_5$	$+$	0	$+$	0	$+$	0	$=$	S_{obs}	$+$	R_S
$x_1 O_{2,1}$	$+\cdots+$	$x_5 O_{2,5}$	$-$	$r_{O/P}\Delta P$	$+$	0	$+$	0	$=$	$O_{2,obs}$	$+$	R_{O_2}
$x_1 PO_{4,1}$	$+\cdots+$	$x_5 PO_{4,5}$	$+$	$r_{P/P}\Delta P$	$+$	$r_D \Delta N^{deni}$	$+$	0	$=$	$PO_{4,obs}$	$+$	R_{PO_4}
$x_1 NO_{3,1}$	$+\cdots+$	$x_5 NO_{3,5}$	$+$	$r_{N/P}\Delta P$	$-$	ΔN^{deni}	$+$	0	$=$	$NO_{3,obs}$	$+$	R_{NO_3}
$x_1 Si_1$	$+\cdots+$	$x_5 Si_5$	$+$	$r_{Si/P}\Delta P$	$+$	0	$+$	0	$=$	Si_{obs}	$+$	R_{Si}
$x_1 DIC_1$	$+\cdots+$	$x_5 DIC_5$	$+$	$r_{C/P}\Delta P$	$+$	$r_D r_{C/P} \Delta N^{deni}$	$+$	ΔC_{inorg}	$=$	DIC_{obs}	$+$	R_{DIC}
$x_1 TALK_1$	$+\cdots+$	$x_5 TALK_5$	$-$	$r_{N/P}\Delta P$	$+$	ΔN^{deni}	$+$	$2\Delta C_{inorg}$	$=$	$TALK_{obs}$	$+$	R_{TALK}
x_1	$+\cdots+$	x_5	$+$	0	$+$	0	$+$	0	$=$	1	$+$	R_Σ

On the left hand side, x_i are in this case the five source water type fractions to compose each observational data through mixing of the source water types (SWT) T_i, S_i, $O_{2,i}$, ... The amount of phosphate added from remineralization of organic matter (ΔP) is linked via a set of first guess Redfield ratios $r_{tracer/P}$ to access the changes in the remaining biogeochemical tracers caused by POM degradation. In the biogeochemical part of the model (columns 6 to 8), the amount of nitrate removed by denitrification is included as ΔN^{deni}. Here r_D is the ratio of phosphorus remineralised for a given loss of nitrate assumed to be 0.01 using the reaction equation of denitrification from [1]. ΔC_{inorg} is the amount of carbon added by the dissolution of calcareous material. The right hand side of the equations is given by the observational data T_{obs}, S_{obs}, $O_{2,obs}$, ... and the residuals (R_{tracer}) occurring from the solution procedure. The last row is an expression of the mass conservation of the mixing components ($\sum x_i = 1$).

If one has nutrient data only and want to consider remineralization processes only (denitrification is of minor importance), one can renounce of using column 7 and 8.

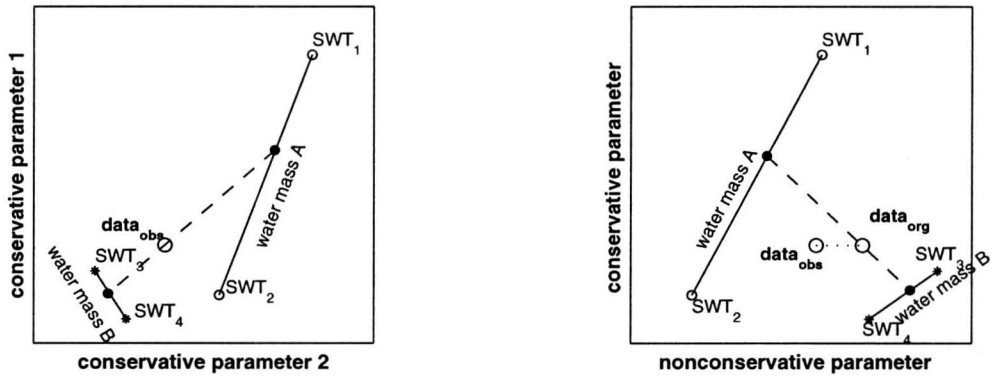

Fig. 3.1: A sketch of mixing in the oceanic thermocline displayed in different parameter subspaces: (left) in a parameter subspace of two conservative parameters and (right) in a parameter subspace defined by one conservative and one nonconservative parameter.

The example of Figure 3.1 is based on the situation usually found in the oceanic thermocline were water masses can be defined through quasi-linear relationships between parameters. The mixing process between two water masses is then formally given by contributions from four source water types ($SWT_{1\text{ to }4}$). All points on the line $SWT_{1/2}$ as well as all points on the line $SWT_{3/4}$ form a water mass (A and B). The final SWT contributions are found by a simultaneous solution in all parameter spaces [6]. Figure 3.1 demonstrates how the extended OMP analysis proceeds in the case of two subspace combinations of conservative and nonconservative parameters. For two conservative parameters (left diagram), OMP analysis

determines the contributions from all four SWT directly. These contributions can be seen as (and if necessary recalculated into) two source water types marked as solid circles on the parameter definition curves of A and B. The observational data ($data_{obs}$) are the result of mixing between water masses A and B along the dashed line. For a nonconservative parameter (right diagram) the observational $data_{obs}$ are also affected by biogeochemical processes. Introducing the Redfield ratio into the linear system separates these effects from the conservative mixing solution, replacing the observational $data_{obs}$ by the corrected values $data_{org}$.

We normalise the elements of G_{ext} in the extended OMP analysis in two stages:
(1) the elements in G without the additional column are normalised as described in section 3.2.
(2) the first guess Redfield ratios are normalised by relating them to the parameter/normalised-parameter range:

$$r_{para/P}{}' = \frac{(G_j)_{max} - (G_j)_{min}}{(G'_j)_{max} - (G'_j)_{min}} \cdot r_{para/P}$$

Here, the normalised ratio ($r_{para/P}{}'$) is obtained by relating the SWT values for j-th parameter of the classical OMP analysis (G_j) to the normalised classical SWT values (G'_j) multiplied by the first guess Redfield ratio ($r_{para/P}$).

4. ANALYSING DATA WITH THE OMP ANALYSIS PACKAGE

4.1 The Package Files

To use the files in this package you should have installed MATLAB Version 4.2 on your machine. To use the GUI version you need MATLAB 5. There are three archives: OMP.zip (4.2 version), OMP_GUI.zip (GUI version) and readme.zip (README.ps and README.html version).

4.2 Version: To set up the OMP analysis, unzip OMP.zip in a directory (eg. /matlab/toolbox/OMP) with:

unip OMP.zip

and add these directory to your MATLAB path.

Now you should have some files in the directory:

- Files you should edit:
  - **OMP_MAIN.M:** The main script, settings has to be edit all marked with a number followed by '!!!!!!!!!!!!!'.
  - **QWT_STEP.M:** A summary file for source water mass definitions.
- Files you should NOT necessarily edit:
  - **NORM_QWT.M:** Normalized the source water types.
  - **CONTOUR_PARA.M:** Example for a contour routine.
  - **QWT_TEST.M:** Shows your SWT relative to data.
  - **DEF_SOURCES.M:** Example file show a way to define the SWT. characteristics
  - **TESTDATA.MAT:** Data from the south-eastern Indian Ocean to test the package.
  - **NANSUM.M:** MATLAB script from P.M. Morgan, calculating the sum but ignoring "NAN".

Additionally (and useful anyway) is the CSIRO seawater library [9] available via anonymous ftp to ftp.ml.csiro.au (192.67.12.100). The files can be found in

/pub/morgan/seawater

as well as the instructions how to set-up your system.

4.2 Analysing Data

To demonstrate the analysis of data with the OMP analysis package, data from the eastern Indian Ocean is analysed step by step. We assume that no biogeochemical changes in the source water characteristics occur, and so the classical OMP analysis is applicable.

The hydrographic structure in this region is dominated by two water masses, the Indian Central Water (ICW) and the Indonesian Throughflow Water or Australasian Mediterranean Water (AAMW), which meet and mix here. We therefore assume to see an intense frontal structure in the data. A simple salinity section will show us this immediately but we are (for some reason) interested in the fractions, say to resolve the effects of mixing on a CFC distribution. Both water masses cover a wide depth (temperature) range of the permanent thermocline and thus must be represented by 4 source water types in our OMP analysis.

4.2.1 Data format and Names of Variables

At first, one has to bring his data in an appropriate format. The easiest way to do this is by using the same variable names that I used and which are listed below:

ptemp	potential temperature
temp	temperature
sal	salinity
pdens	potential density
oxy	oxygen
ni	nitrate
ph	phosphate
press	pressure
comment	comment on the data (region, origin, ...)
lat	latitude of stations
long	longitude of the stations
date	date of the cruise

Additional parameters can be added with additional names, e.g. "si" for silicate. How to transfer the data into MATLAB *.MAT format: This is archived by saving the data from the MATLAB workspace with the command 'save filename'. This will made all variables (including the variables names) available to us, just with the 'load filename' command. Parameter gaps should be filled with a dummy value of "-9".

If one likes to use additional parameters (e.g. silicate), they must be included in the **OMP_MAIN.M** script (at the places numbered with #4., #9., and #10. see section 4.2.4 for details) as well as in the source water definition script **QWT_STEP.M** (explained in the next section).

4.2.2 Define the Source Water Types

To define the source water masses values (SWT), literature values can be used or, what is usually a better way, one defines the values from data in the source region. To do so, plot all parameters against temperature for data in the source region of the water masses (Figure 4.1). This is done under the assumption that temperature is the important stratification parameter (in polar regions, this shouldn't be true!).

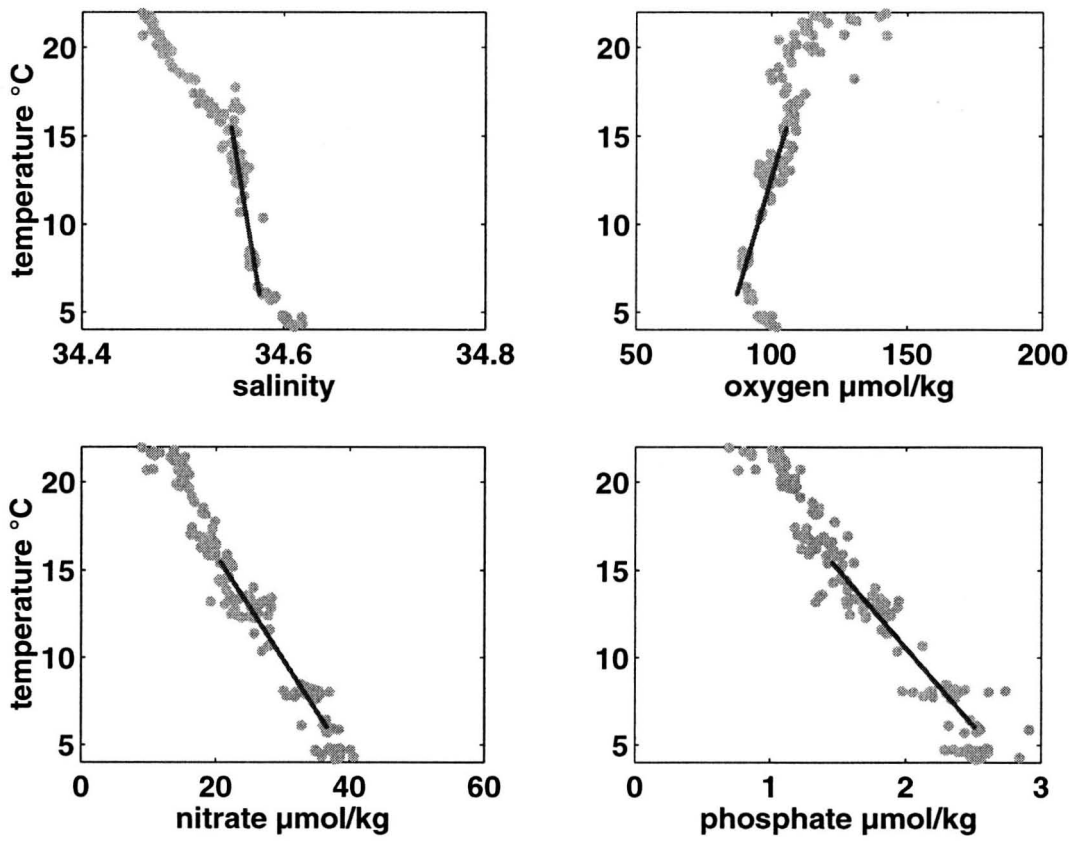

Fig. 4.1: Plot of all parameters used in the OMP analysis example relative to temperature (•). The lines indicate the linear fit obtained from the MATLAB POLYFIT.M script.

In our example we have data from one source region only available (the AAMW source region) and we took data from the literature for the second source values (for the Indian Central Water source). The water mass described by the available source data, should have the largest variability which is important to calculate the weights. (See file **DEF_SOURCES.M** as an example how to do this!)

Results from linear regression in source region of AAMW			
Parameter	upper SWT	lower SWT	mean(δ_j)
Temperature °C	15.443	9.4	-
Salinity	34.548	34.69	0.0046
O_2 μmol kg^{-1}	105.4	87.5	2.85
PO_4 μmol kg^{-1}	1.5	2.6	0.149
NO_3 μmol kg^{-1}	20.7	36.6	2.037

Tab. 4.1: Parameter values obtained in source region of one water mass

To determine the temperature range with approximately linear relationships between **all** parameters, we used the MATLAB script POLYFIT.M. It delivers (with the script POLY-

VAL.M) an error estimate of the fit. This error can be useful in detecting the upper and lower limit of linear parameter behaviour as well as to calculate the weights for the OMP analysis (δ_j; see section 3.2).

Taking the end points of the line (Figure 4.1) as the source water types G_{ji}. Mixing along the SWT lines (lines in Figure 4.1) and the accompanied variability (δ_j) represent mathematically all parameter points of a physical water mass.

Now we are ready for the first OMP analysis setting:
Now one has to add the so obtained SWT values (table 4.1) to the script **QWT_STEP.M** for later use with **OMP_MAIN.M** (same for GUI users).

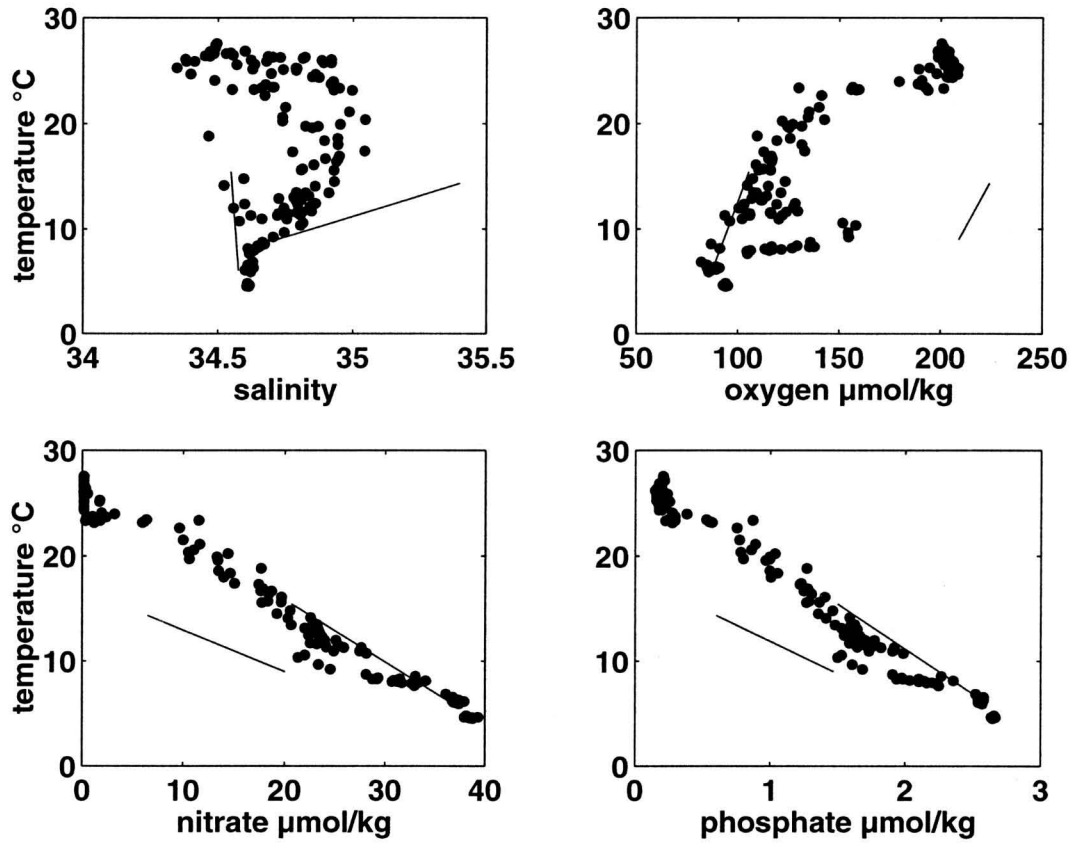

Fig. 4.2: Testdata set (•) and chosen source water masses.

The final definitions of both water masses, represented by four SWT values can be seen in figure 4.2.

4.2.3 Determine the Weights

Next step is to determine the weighting matrix:
We use in our example the mean variance of the linear fit to the data in the source region, related to the variance in the source water type matrix (calculated with the 'stdG' output from **NORM_QWT.M**) to determine the weighting matrix (see section 3.2 for details). Col-

umn 4 in table 4.1 summarises the so calculated values. With the weighting formula (**see section 3.2 for details**) we obtain the weights for each parameter while temperature, and mass conservation are weighted as salinity (highest weights).

Now we are ready for the second OMP analysis setting:
Put the weights in the script **OMP_MAIN.M** (place #5.) or place at in the GUI at 4.

Parameter	Weight
Temperature	87
Salinity	87
O_2	25
PO_4	6
NO_3	6
mass conservation	87

Tab. 4.2: Parameter weights

If one uses single source water types to represent a water mass, the variability around that point in the parameter space has to be used to calculate the weights.

4.2.4 Edit OMP_MAIN.M and Run it

Now edit the **OMP_MAIN.M** script. If necessary, change the lines (marked with # !!!!!!!!!!!!!):

ext 1: Decide if you want the "extended" or the "classical" OMP (uncomment it).

1. : Enter as string the name of the dataset you want to analyse.

2. : Enter as string the name of the water masses you want to analyse. If you don't know the name, type something like 'w1';'w2';'w3';... in the brackets.

3. : Enter in the find command the restrictions on the parameters - using only datapoints where all biogeochemical parameters and oxygen are present is one possibility.

4. : Check the next block for all data names involved in your analysis. They will are cutted to the 'index' interval. (I don't know why there is these transpose operator !).

5. : Add the here weights you obtained from the analysis in the source region.

ext 2: Insert the Redfield ratio you want to use for the analysis. Be careful to choose the same parameter order then in the source water matrix !!.

6. : Add here the position of the SWT definition in your **QWT_POS.M** file.

7. : Add again the names of the water masses - but not as string (without ' ').

8. : Who sum up the SWT to water masses. In our example, we have four SWT representing two water masses. Water mass 1 is represented by SWT 1 and 2 and is called ICW, water mass 2 is represented by SWT 3 and 4 and is called AAMW. Enter as string variable (watch out for equal length !!).

9. : Check the following block for you variables included in the OMP analysis. Let's say you want to include silicate as well: add si_dat = silicate(k); and add this as well in the vector btst !!!!

10.: Check if the btst vector includes all variables from your OMP analysis and additionally if these values are all at the place corresponding to the **QWT_STEP.M** order.

Now all should be setup for analysing. Run the **OMP_MAIN.M** script:
The visible output ('figure window') is a plot of the residual distribution from the mass

conservation equation relative to pot. density (figure 4.3). This can be seen as a quality parameter for the fit.

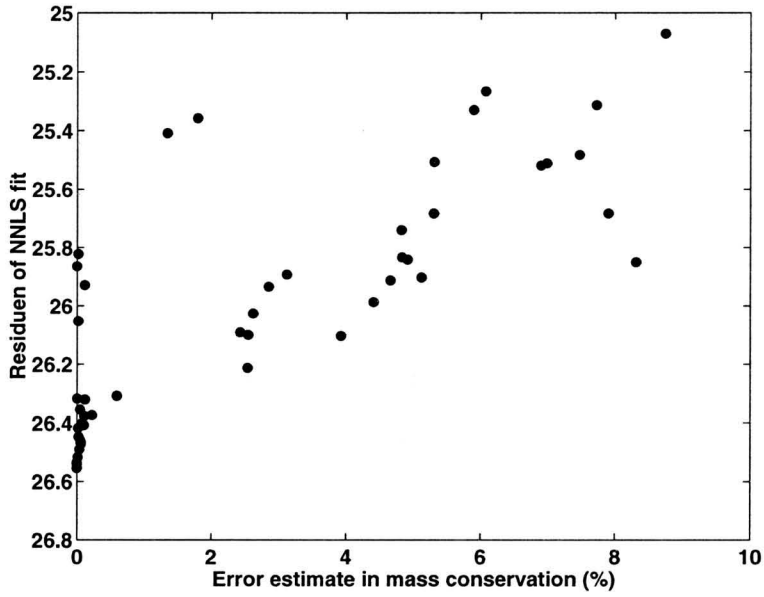

Fig. 4.3: Estimated error of the mass conservation equation from **OMP_MAIN.M** in percent. The increase in error towards the surface may reflect biogeochemical cycling.

To visualise the results, the **CONTOUR_PARA.M** script is calculating distances between stations of a section (here sorted by latitude) and grid a parameter (here the ICW water mass content) on a regular grid and finally contours it (Figure 4.4). The front between the two water masses is clearly visible and is now quantified by the amount (percent) of the SWT.

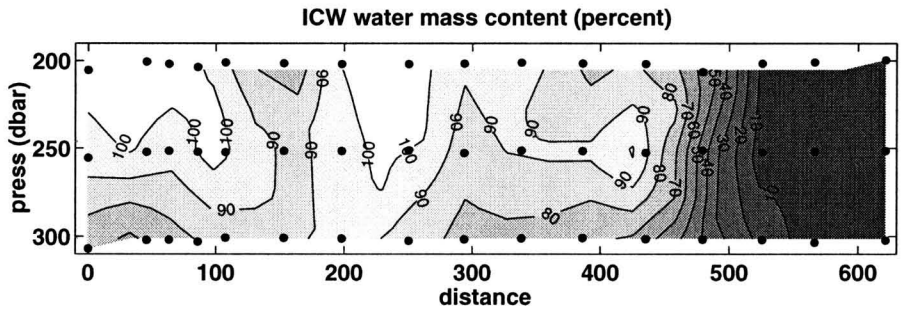

Fig. 4.4: Amount if the ICW component (percent) of the testdata set.

If you look into the MATLAB workspace (type 'whos' at the prompt), it will show up with a lot of variables. Some may be interesting for further interpretation of the OMP analysis results:

A	Result matrix, columns are analysed data points, rows water masses content for both water masses involved in analysis.
AAMW	Water mass content of AAMW component.
Dual	Fitting parameter from NNLS.M script [8].
G1	Source water matrix (from QWT_STEP.M).
ICW	Water mass content of ICW component.
Wx	Weighting matrix.
dataset	Name of the dataset.
quali	Quality of the fit (residues of mass conservation equation).
a	Quantity describing the changes in the parameter field due to biogeochemical cycling (only in case of the "extended" OMP).
ratio	first guess Redfield ratio.

4.2.5 Improving the Results

The quality of the results may be improved by changing the SWT values input characteristics in the **QWT_STEP.M** script.
Trying other SWT values for the ICW component:
Change in the **OMP_MAIN.M** script at #6. the position of the SWT from [1 2 3 4] to [1 2 5 6] and the script will take row 1, 2, 5, and 6 from QWT_STEP.M as the new input. Run OMP_MAIN again, and the results will be changed to lower residuals of the fit, hence the second definition is a better source to compose the observational data (variable 'quali'; Figure 4.5).

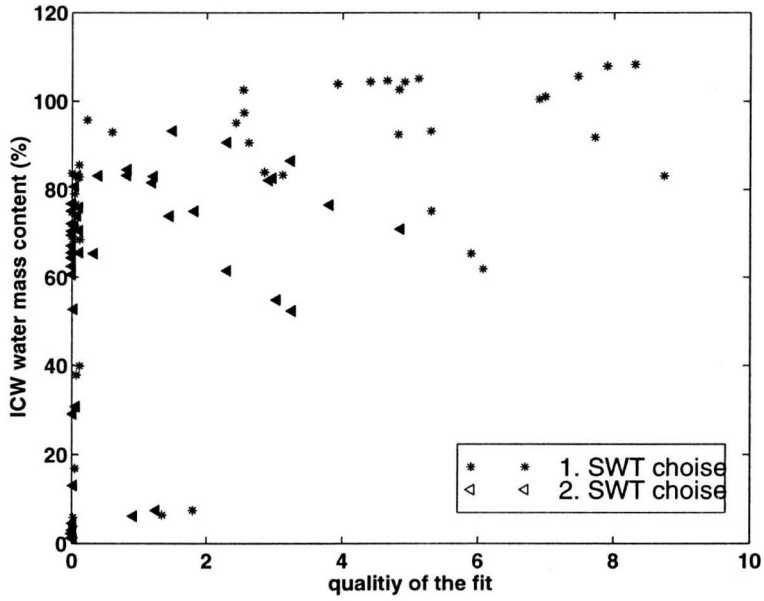

Fig. 4.5: Change in the residuals of the fit (variable 'quali' in MATLAB workspace), through change in SWT input.

The fit is still not excellent, but you may improve it by taking better SWT definitions. It is evident that the error is correlated to the amount of ICW, therefore a change in ICW SWT is useful. The results shown here are obtained with higher weighting on the mass conservation

equation (factor of 287 instead of 87) which is also a point to think about - what is a valid weight for mass conservation ??

4.3 GUI Version

There is a GUI version of the package included (OMP_GUI.zip), which will only run under MATLAB 5.1. You should find the following files after unzipping:

GUI_OMP.m	Start the GUI version
GUI_OMP.mat	Automatically generated file for GUI_OMP.m
modul_load.m	Module to load dataset to GUI
modul_main.m	Module to make OMP analysis
modul_plo.m	Module to plot parameters at GUI
modul_sel.m	Module to select specific data from the whole dataset
omp_dat.mat	Automatically generated file from GUI_OMP.m
	Same files as in OMP.zip:
nansum.m	Script from P. Morgan (CSIRO) to build the sum of matrix columns (ignoring NaN)
norm_qwt.m	File use by modul_main.m
qwt_step.m	Definition values for the SWT values
testdata.mat	Test dataset

To start the GUI just type 'GUI_OMP' at the MATLAB prompt (if you copied the files to a place included in your MATLAB path variable). The interface is automatically generated by MATLAB's **uide.m** and looks like that (ugly!). It calls the **modul_*.m**, files to execute the steps of the OMP analysis.

The switches are in general the same as in the 4.2 Version. Explanation is given with numbers and letter as (figure 4.6):

1	Name of the dataset to analyse (The format is explained in section 4.2.1).
2	Data range to analyse, say "oxy> 100&pdens< 27.2" will force to analyse only data with an oxygen content greater 100 and a potential density smaller 27.2 kg m^{-3}.
3	Positions of the SWT values in the **QWT_STEP.M** file.
4	Weight of the parameter.
5	Display some infos while executing OMP.
6 and 7	Switch for the "extended OMP analysis", if the switch is "on", you must define appropriate first guess Redfield ratios in 7.
9 and 10	Parameters to display in the figure window (8).
11 and 12	Parameter you want to use for the OMP analysis (11), and parameter which are not analysed but used for displaying your results (12; e.g. latitudal position).

There are four push bottoms to activate the analysis:

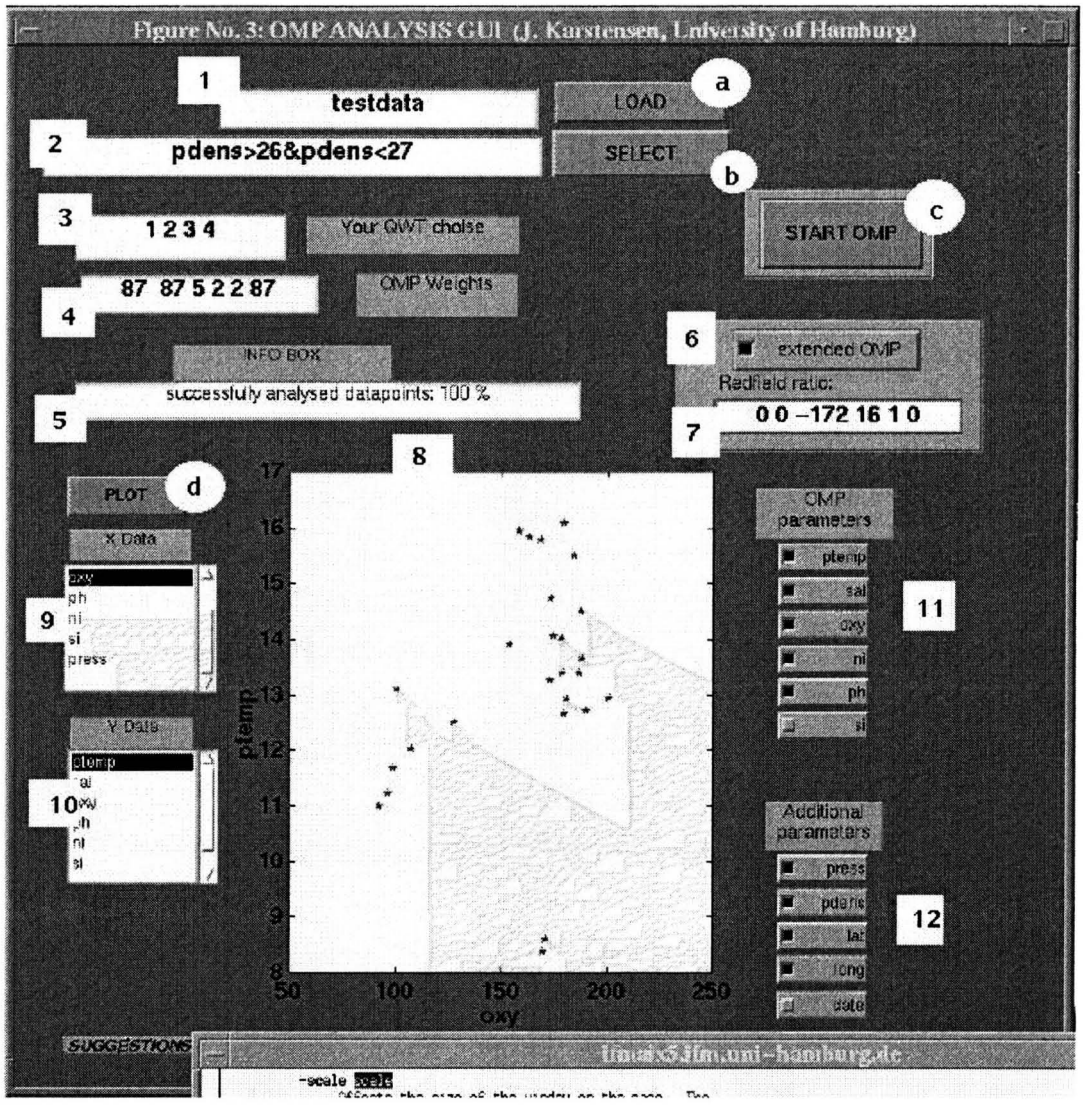

Fig. 4.6: Screen shot of the GUI version. Numbers and letters are explained in the text.

- **a** Load the dataset given in **1**.
- **b** Select the data range given in **2**.
- **c** Plot parameters against each others (selected in 9 and 10).
- **d** Start the OMP analysis (with settings in **1, 2, 3, 4, 8, 11,** *(6, 7 if the extended version is used)*).

After analysing (Push bottom d), MATLAB will pop up a figure showing the residuals of the mass conservation to give you a first impression on the quality of the least-square fit (see section 4.2.5 for details).

BIBLIOGRAPHY

[1] L. A. Anderson and J. L. Sarimento. Redfield ratios of remineralization determined by nutrient data analysis. *Global Biogeochem. Cycles*, 8:65–80, 1994.

[2] W. S. Broecker. 'NO', a conservative water-mass tracer. *Earth Planet. Sci. Lett.*, 23:100–107, 1974.

[3] R. Chester. *Marine Geochemistry.* Chapman & Hall, London, U.K., 698 pp., 1990.

[4] J. P. Jacobsen. Eine graphische Methode zur Bestimmung des Vermischungskoeffizienten im Meer. *Gerlands Beiträge zur Geophysik*, 16:404–412, 1927.

[5] J. Karstensen and M. Tomczak. Age determination of mixed water masses using CFC and oxygen data. *J. Geophys. Res.*, 103(C9):18599–18610, 1998.

[6] C. L. Lawson and R. J. Hanson. *Solving least square problems.* Prentice-Hall, 1974.

[7] D. L. Mackas, K. L. Denman, and A. F. Bennett. Least-square multiple tracer analysis of water mass composition. *J. Geophys. Res.*, 92(C3):2907–2918, 1987.

[8] W. Menke. *Geophysical data analysis, discrete inverse theory.* Academic Press Inc., 1984.

[9] P. P. Morgan. SEAWATER: A library of MATLAB computational routines for the properties of sea water. Technical Report 222, CSIRO Marine Laboratories, 1994.

[10] R. O. Thompson, R. Y. Edwards, and R. J. Edwards. Mixing and water mass formation in the Australian Sub-Antarctic. *J. Phys. Oceanogr.*, 11:1399–1406, 1981.

[11] M. Tomczak. An analysis of mixing in the frontal zone of South and North Atlantic Central Water off North-West Africa. *Progress in Oceanography*, 10(3):172–192, 1981.

[12] M. Tomczak and D. G. B. Large. Optimum multiparameter analysis of mixing in the thermocline of the eastern Indian Ocean. *J. Geophys. Res.*, 94(C11):16141–16149, 1989.

[13] G. Wüst. *Wissenschaftliche Ergebnisse der Deutschen Atlantischen Expedition auf dem Forschungs- und Vermessungsschiff "Meteor" 1925-27, Band 6, erster Teil, Ozeanographie.* A. Defant, 1935.

[14] Y. You. Dianeutral mixing in the thermocline circulation of the Indian Ocean. *Deep-Sea Res.*, 43 (I):291–320, 1996.

[15] Y. You and M. Tomczak. Thermocline circulation and ventilation in the Indian Ocean derived from water mass analysis. *Deep-Sea Res.*, 40:13–56, 1993.

APPENDIX

A. SOME OF THE ORIGINAL FILES

The original **OMP_MAIN.M**:

```
% ######### OMP analysis main program  ###################
%
% This is a first try on a easy-to-handle package to use the
% OMP analysis to resolve the fractions of water masses involved
% in the mixing at a observation point in the ocean.
% Some preparation work has to be done before applying data to
% this program (see README.ps or REDAME.html for more details):
%
% 1. produce a *.MAT file containing all nessesary variables
%    (for names see README.html)
% 2. change all rows marked with %  !!!!!!!!!!!! to your personal settings
%    (depending on the water masses; region; etc.)
%
%
%
% CALL: qwt_step.m qwt_tst.m nansum.m (Philip Morgan, CSIRO)
%
%
% This program is part of the OMP package from:
% Institut fuer Meereskunde
% J. Karstensen
% Troplowitzstr. 7
% 22529 Hamburg
% Germany
%
% BUGS: karstens@ifm.uni-hamburg.de

clear all


% switch if you want classic or extended OMP (See README.PS for details)
%OMP='ext';   % ext. 1 !!!!!!!!!!!!!!!!!!!
OMP='cla';


% ***********************************
% dataset *.mat file name:
dataset='testdata';  % 1. !!!!!!!!!!!!


%***********************************
% name of the resulting water masses:
```

```
WM_stri=str2mat([ 'AAMW'; ' ICW']);   % 2. !!!!!!!!!!!!


disp(['      dataset:  ' dataset ' '])
eval(['load ' dataset])

%************************************
% sort out data through specific criteria:
index=find(oxy>=0&ph>=0&ni>=0&press>180&press<310);  % 3. !!!!!!!!!!!!!!

% cutting the data
% (CHECK FOR YOUR PARAMETERS !!) % 4. !!!!!!!!!!!!!!!!!!!!!
      lat=  lat(index)';
    press=press(index)';
     long= long(index)';
     temp= temp(index)';
    ptemp=ptemp(index)';
      sal=  sal(index)';
      oxy=  oxy(index)';
       ph=   ph(index)';
       ni=   ni(index)';
    pdens=pdens(index)';
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
clear index
save omp_dat


% number of water masses
nr_of_wm = size(WM_stri,1);
 disp('   ')
 disp([' water masses involved: ' num2str(nr_of_wm)])
 disp('   ')

for i=1:size(WM_stri,1)
  eval([WM_stri(i,:) '=' num2str(i) ';']);
end

%************************************
% Weigthing Matrix (see manual for details on calculating this)
Wx=diag([87 87 25 6 6 87]); % 5. !!!!!!!!!!!!!!!

if OMP(1:3)=='ext'
% Redfield ratio
  ratio=[ 0. 0. -1.7  0   .01   .16   0.];   % ext. 2 !!!!!!!!!!!!!!!
end


% some settings for storage (in the first step only)
  A    = zeros(nr_of_wm,length(press))-nan;


% load dataset (standard)
load omp_dat  % do NOT change this!!
```

```
%************************************************
% position (row) of source water types in QWT_STEP.M
      qwt_pos = [ 1 2 3 4 ];                % 6. !!!!!!!!!!!!!!

G1=qwt_step(qwt_pos);

%************************************************
% again name of the water masses involved in this analysis
  wm_index    = [AAMW ICW  ];                % 7. !!!!!!!!!!!!!!


%************************************************
% Who sum up the calculated fractions of the water masses involved:
 store_string = ['1:2';'3:4' ];           % 8. !!!!!!!!!!!!!!

  wm = length(wm_index);
[m,n]=size(G1); % number of water types

Wx(m,m)=max(max(Wx)); % maximum weight to mass conservation

% normalise the source water matrix (get meanG, get stdG for weighting):
[G,mG,stdG]=norm_qwt(G1);


% EXTENDED OMP switch:
if OMP(1:3)=='ext'
   % Adding Redfield ratio to the system:
 G1(1:m,n+1)=[ratio(1:m)]';
   % 2. normalisation of the ratios:
   %--------------------------------
 for rr=1:(m-1)
  G(rr,n+1)=ratio(rr)*(max(G(rr,1:n))-min(G(rr,1:n)))...
           /(max(G1(rr,1:n))-min(G1(rr,1:n)));
 end
 G(m,n+1)=0;
end


% adding weights
G2=Wx*G;


gap=0;

%************************************************************
% Vector of each datapoint (btst) is build here

for k=1:length(lat);

% (CHECK FOR YOUR PARAMETERS !!) 9. !!!!!!!!!!!!!!!!!!!
     p_dat = press(k);
     t_dat = ptemp(k);
     s_dat =   sal(k);
```

```
     oxy_dat =   oxy(k);
     ph_dat =    ph(k);
     ni_dat =    ni(k);
     pden_dat = pdens(k);
     kon=1;


btst= [t_dat,s_dat,oxy_dat,ph_dat,ni_dat,kon]; %10. !!!!!!!!!!!!!!


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%% leave ALL UNCHANGED beyound this point !!  %%%%%%%%%%%%%%%%%

%looking for GAP:
   index0=find(btst < -1);
   index1=find(btst > -1);

cutit=n;
% using extended OMP we need one parameter more
%      (because we have one unknown more!)
if OMP(1:3)=='ext'
  cutit=n+1;
end

if length(index1) < cutit
% not enough parameters to find a NNLS fit
% DATA point not successful analysed
disp(['ANALYSIS of the datapoint failed, not enough parameters available !!'])

    A(wm_index(all_wm),(k)) = nan;
 Dual(wm_index(all_wm),(k)) = nan;
 gap=gap+1;

else
  %new data without GAP:
   b1 = btst(index1);
   mG1=   mG(index1);
 stdG1= stdG(index1);

  % standardize the data:
   for i=1:length(b1)-1
     b(i,1)=(b1(i)-mG1(i))/stdG1(i);
   end
     b(length(b1))=b1(length(b1));

  %add weights:
     b2=Wx(index1,index1)*b;

  %optimization algorithm
 [x,dual] = nnls(G2(index1,:),b2);

for all_wm=1:wm
    A(wm_index(all_wm),k) = sum(x   (str2num(store_string(all_wm,:))));
```

```
 Dual(wm_index(all_wm),k) = sum(dual(str2num(store_string(all_wm,:))));
  eval([WM_stri(all_wm,:) '(k)=' num2str( A(wm_index(all_wm),k)) ';']);
end

% in case of extended OMP store biogeochemical part:
if OMP(1:3)=='ext'
  a(k)=x(length(x));
end


   clear b
   end % of last IF
 end

if gap>0
disp(['successfully analysed datapoints:' k*100/gap ' %' ])
 disp('   ')
 disp('   ')
else
disp(['successfully analysed datapoints: 100 %' ])
 disp('   ')
 disp('   ')
end

% plotting residuals
quali=(nansum(A(1:nr_of_wm,:))-1)*100;
figure
plot(quali,pdens,'.','markers',20),axis('ij')
ylabel('Residuen of NNLS fit')

clear dual WM_stri wm_index k gap x index0 index1 p_dat ni_dat t_dat s_dat ...
    ph_dat wm store_string i n in kon m mG mG1 b1 b2 btst stdG stdG1 qwt_pos ...
    nr_of_wm pden_dat G G2  all_wm oxy_dat
```

The original **QWT_STEP.M**:

```
function G1=qwt_step(wm_row)
%  This is a summary of all water mass definition
%  by sending the wm_row you get out the G1 matrix (source water matrix)
%  used by the OMP analysis
%
%  To find the right rows: have a look at this file
%
% should be a part of the OMP packet
%
%1  upper AAMW
%2  lower AAMW
%3  upper ICW ***** first set
%4  lower ICW ***** first set
%5  upper ICW ***** second set
```

```
%6  lower ICW ***** second set
%
%
%
% This program is part of the OMP package from:
% Institut fuer Meereskunde
% J. Karstensen
% Troplowitzstr. 7
% 22529 Hamburg
% Germany
%
% BUGS: karstens@ifm.uni-hamburg.de

if nargin<1
 disp(' ')
 disp(['     Please give the rwo index of SWT definitions you want to use:'])
 disp([' e.g. qwt_step([1 2 3 6]) activates row 1, 2, 3, and 6 '])
 disp('  '),return

end

all=[...

% AAMW
%ptemp  sal oxy PO4 NO3 mass
6.016 34.576 87.4 2.6 36.6 1.;... %1 lower AAMW
15.443  34.548 105.4 1.5 20.7 1.;... %2 upper AAMW

%ICW
%ptemp  sal oxy PO4 NO3 mass
9 34.72 209 1.47 20 1.;... %3 lower ICW
14.35 35.4 224 0.6 6.5 1.;... %4 upper ICW

%another ICW definition
%ptemp  sal oxy PO4 NO3 mass
9.5 34.75 265 0.88 10.7 1.;... %5
16.53 35.668 218 0.32 2.4 1.;... %6

];

G1=all(wm_row,:)';
```

The original **QWT_TEST.M**:

```
function qwt_test(para1,para2,wm_row);
%   shows the SWT definition relative to data
%     !!  only if standard QWT order: !!
%
%  eg.  T/S check with water type 2,4,6,8 is
%    >> qwt_tst(2,1,[2 4 6 8])
%
```

```
%  parameter 'para1(x-axis)/para2(y-axis)' coding:
%    1 = ptemp
%    2 = salinity
%    3 = oxygen
%    4 = silicate
%    5 = phosphate
%    6 = nitrate
%
%  CALL: QWT_STEP.M
%
% This program is part of the OMP package from:
% Institut fuer Meereskunde
% J. Karstensen
% Troplowitzstr. 7
% 22529 Hamburg
% Germany
%
% BUGS: karstens@ifm.uni-hamburg.de

if nargin<3,disp('define wm_row for QWT_STEP.M !'),break,end
if nargin<2,disp('secound parameter !'),break,end

G1=qwt_step(wm_row);

hold on
plot(G1(para1,:),G1(para2,:),'o')

for k=1:length(wm_row)
 text(G1(para1,k),G1(para2,k),num2str(wm_row(k)), ...
 'fontsize',14,'horizontalal','right','fontweight','bold')
end

return
```

The original **NORM_QWT.M**:

```
function [G,mG,stdG]=norm_qwt(G1)
% Normalises the source water type (SWT) matrix G1
% and calculate standarddeviation and mean
%
%  INPUT:
%   G1 : Input nonnormalized SWt matrix
%
% OUTPUT:
% G : normalized SWT matrix
% mG : mean original SWT matrix
% stdG : standrddeviation of original SWT matrix
%
% called by OMP_MAIN.M
%
%
% This program is part of the OMP package from:
% Institut fuer Meereskunde
% J. Karstensen
% Troplowitzstr. 7
% 22529 Hamburg
% Germany
%
% BUGS: karstens@ifm.uni-hamburg.de


[m,n]=size(G1); % number of water types

% mean and standarddeviation of SWT
  mG = mean(G1');
stdG =  std(G1');

% standardize QWT for m eq. 7
for i=1:n
  for kkk=1:m-1
   G(kkk,i)=(G1(kkk,i)-mG(kkk))/stdG(kkk);
  end
   G(m,i)=G1(m,i);  % mass untouched
end
```